兰州大学学报(自然科学版), 1998, 34(1): 25~ 29
JOURNAL OF LANZHOU UNIVERSITY(Natural Sciences)

数据通信中一类信息打包问题的最优解 ※

李效虎　李泽慧　　　　　　陈元清
(兰州大学数学系, 730000　兰州)　(广东发展银行科技部, 510120　广州)

摘　要　对于数据通信过程中存在的一类信息打包问题建立了概率模型,并就打包的方案进行讨论,通过定量分析获得了最佳打包方案的存在性,并给出了几个数值模拟结果.

关键词　打包　Poission过程　数学期望
中图法分类号　O212.2

0　引言

在数据通信过程中,往往需要将一个完整的源信息群分解成一些小的片段,然后给这些小片段的信息再附加上其它信息后进行传输,以防止在信息传输过程中出现失真、泄密等.以上的工作通常被称为打包(Packing).打包能够增加数据传输的安全性和可靠性,但打包由于需要附加额外信息而导致实际传输的信息长度比源信息的长度大,势必增加传输费用和传输时间,作为管理人员,当然希望传输费用不致于太大,使数据传输经济可行;所以研究传输过程以获得一个最佳的打包方案是十分必要的.

首先考虑打包的工作环境:

A　作为信息的输入端,由于昂贵的外设,往往是由许多部门(信息源)所共享,所以源信息的到达规律是一个随机现象,而且每次到达的源信息长度也是不确定的.为此,假定源信息的到达服从一个局部有限的随机点过程;而且每次到达的源信息长度 Y_i 同分布,到达间隔 X_i 满足: (X_i, Y_i), $i = 1, 2, \cdots$ 相互独立.考虑到源信息有许多产生源,进一步假定到达过程与信息长度大小相互独立.

B　打包时将源信息分解为长度为 l 的信息片段,不足 l 长度的均按 l 长度计算,这样除传输一些附加信息外还要传输一些空白信息,传输方也要为此付费,假设每一个长度为 l 的信息包的传输费用为 $C_0(l)$.

C　为经济起见,传输方除考虑传输总费用不致太大外,也希望在传输过程中无用信息的传输费用尽可能地小一些.由于无用信息中的附加信息的长度决定于通信技术,其长度是不可删减的,所以我们只考虑空白信息所生产的额外传输费用.

1　概率模型的建立

以 $N(t)$ 记在时间间隔 $(0, t]$ 中到达的源信息段数,由于到达过程局部有限,所以对 $0 < t$

※ 收稿日期: 1996-12-04.
李效虎, 29岁, 男, 讲师.

$<\infty, EN(t) < \infty$.若$\{N(t), t \geqslant 0\}$服从强度为λ的 Poisson过程,则对任一 $t > 0$, $N(t)$具有参数为λt的 Poisson分布[1].

记
$$\{x\} = \begin{cases} [x+1], & \text{当 } x > [x], \\ [x], & \text{当 } x = [x]. \end{cases}$$
其中: $[x]$表示不超过 x的最大整数.这样$(0, t]$间隔中到达的源信息打包传输的总费用为

$$C_T(t) = C_0(l)\sum_{i=1}^{N(t)}\{\frac{Y_i}{l}\}. \tag{1}$$

利用著名的 Wald方程[1]可以得到平均总费用

$$EC_T(t) = C_0(l)EN(t)E\{\frac{Y_1}{l}\} = \lambda t C_0(l)E\{\frac{y_1}{l}\} \quad (N(t)\text{为 Poisson过程}). \tag{2}$$

若 Y_1具有参数为$_^{-1}$的指数分布[2],则

$$\begin{aligned} E\{\frac{Y_1}{l}\} &= \sum_{k=1}^{\infty} kP((k-1)l < Y_1 \leqslant kl) \\ &= \sum_{k=1}^{\infty} k(e^{-\_(k-1)l} - e^{-\_kl}) \\ &= e^{l}/(e^{l} - 1). \end{aligned} \tag{3}$$

容易验证(3)关于 l递减.考虑到不论 Y_1的具体形式$\{\frac{Y_1}{l}\}$关于 l递减, $E\{\frac{Y_1}{l}\}$必关于 l不增.

间隔$(0, t]$中被打包的空白信息的传输总费用为

$$C_W(t) = \frac{C_0(l)}{l}\sum_{i=1}^{N(t)}(l\{\frac{Y_i}{l}\} - Y_i). \tag{4}$$

其平均值为

$$\begin{aligned} EC_W(t) &= \frac{C_0(l)}{l}EN(t)E(l\{\frac{Y_1}{l}\} - Y_1) \\ &= \lambda t C_0(l)(E\{\frac{Y_1}{l}\} - E(\frac{Y_1}{l}))\ (N(t)\text{为 Poisson过程}). \end{aligned} \tag{5}$$

若 Y_i具有参数为$_^{-1}$的指数分布,则

$$E\{\frac{Y_1}{l}\} - E(\frac{Y_1}{l}) = e^{l}/(e^{l} - 1) - 1/\_l. \tag{6}$$

利用(2)和(5),间隔$(0, t]$中空白信息的费用与总费用的比率为

$$\begin{aligned} EC_W(t)/EC_T(t) &= E(l\{\frac{Y_1}{l}\} - Y_1)/lE\{\frac{Y_1}{l}\} \\ &= 1 - EY_1/lE\{\frac{Y_1}{l}\}. \end{aligned} \tag{7}$$

当 Y_1具有参数为$_^{-1}$的指数分布时,(7)成为

$$EC_W(t)/EC_T(t) = 1 - (e^{l} - 1)/\_le^{l}. \tag{8}$$

不难验证,(8)关于 $l > 0$递增.

根据对传输的总体要求,所谓最佳打包方案就是指寻求合适的打包长度 l,使传输总费用 $EC_T(t)$不必太大(小于等于给定的常数 $C > 0$)的同时,也使空白信息的传输费用在总费用中的所占的比例 $EC_W(t)/EC_T(t)$不至于过大(小于等于给定的 $T > 0$).在下一部分中,我们将讨论模型解的存在性.

2　模型解的存在性

若数据传输部门能够接受的单位时间上的最大平均总传输费用为 $C > 0$,允许的单位时间上空白信息的传输费用与总费用的最大比例系数为 $0 < \mathrm{T} < 1$,则模型的解应是满足以下两式的一些 l

$$EC_T(t)/t = C_0(l)EN(t)E\{\frac{Y_1}{l}\} \leqslant C, \tag{9}$$

$$EC_W(t)/EC_T(t) = 1 - EY_1/lE\{\frac{Y_1}{l}\} \leqslant \mathrm{T}, \tag{10}$$

以上两式分别来自(2)和(7).(9)依赖于时间参数 t,我们代之以考虑在极限意义下的平稳平均费用,利用基本更新定理[1,3]知

$$EC_T(t)/t = C_0(l)E\{\frac{Y_1}{l}\}EN(t)/t \to C_0(l)E\{\frac{Y_1}{l}\}/EX_1, \quad t \to \infty. \tag{11}$$

其中 X_1是首次源信息到达时间间隔.若 Y_1具有参数为 _ $^{-1}$的指数分布而源信息到达服从强度为 λ 的 Poisson过程,利用(9),(11)可以变为

$$EC_T(t)/t \to C_0(l)\lambda e^{l}/(e^{l} - 1), t \to \infty. \tag{12}$$

注意,这时 $\{C_T(t), t \geqslant 0\}$实质上是一个复合 Poisson过程.这样便有如下结论:

定理 1　对单位时间上平均总传输费用及空白信息传输费用占总传输费用比例带约束的模型的解由不等式组(9),(10)确定.

以下讨论不等式组(9),(10)解的存在性.

引理 2　函数 $l\{\frac{x}{l}\}$关于 $l > 0$具有局部单调递增性,其中 $x > 0$确定,$\{\frac{x}{l}\}$由(1)确定.

证明　当 $l \geqslant x$时,$l\{\frac{x}{l}\} = l$;当$\frac{x}{k} \leqslant l \leqslant \frac{x}{k-1}$时($k > 1$为正整数),$k - 1 < \frac{x}{l} \leqslant k$.设$\frac{x}{k} \leqslant l_1 < l_2 < \frac{x}{k-1}$,则

$$l_2\{\frac{x}{l_2}\} - l_1\{\frac{x}{l_1}\} = \begin{cases} k(l_2 - l_1) + 1 > 0, & l_1 > \frac{x}{k}, \\ k(l_2 - l_1) > 0, & l_1 = \frac{x}{k}. \end{cases}$$

可见,$l\{\frac{x}{l}\}$在$[\frac{x}{k}, \frac{x}{k-1})$上严格递增.进一步还有

$$l\{\frac{x}{l}\} \to \infty, \quad l \to \infty; \qquad l\{\frac{x}{l}\} \to \frac{x}{k} + x, \quad l \to \frac{x}{k}.$$

定理 3　对适当的 C及 T,若 $C_0(l) = C_0$,则满足(9)与(10)的 l是存在的.

证明　当 $C_0(l) = C_0$时,(11)成为

$$EC_T(t)/t \to C_0E\{\frac{Y_1}{l}\}/EX_1,$$

上式右侧关于 l单调不增,所以必存在某个 $l_1 > 0$,使 $l \geqslant l_1$时,

$$C_0E\{\frac{Y_1}{l}\}/EX_1 \leqslant C. \tag{13}$$

事实上,若 Y_1 具有参数为 $\_^{-1}$ 的指数分布,则欲使(13)成立,只要 $\frac{C_0}{EX_1}\cdot\frac{e^{l}}{e^{l}-1}\leqslant C$.记 $M=C_0/EX_1$,当 $M>C$ 时,只要 $l\leqslant\frac{1}{\_}\log\frac{C}{M-C}$ 即可;而当 $M\leqslant C$ 时,对一切 $l>0$ 均可.又由引理2,$lE\{\frac{Y_1}{l}\}$ 与 $l\{\frac{x}{l}\}$ 具有相同的局部单调不减性,从而在(10)中

$$1-EY_1/lE\{\frac{Y_1}{l}\}=EC_W(t)/EC_T(t)$$

也具有相应地单调不减性,所以,对合适的 $0<T<1$,必存在集合 $L\subset R$ 使当 $l\in L$ 时,成立

$$EC_W(t)/EC_T(t)\leqslant T. \tag{14}$$

事实上,在 Y_1 具有参数为 $\_^{-1}$ 的指数分布时,对给定 $0<T<1$,令 $W=1-T$,要使(14)成立,只须 $W\_l+e^{-\_l}\leqslant 1$.注意到,由 Taylor公式有 $e^{-\_l}\geqslant 1-\_l+\_^2l^2/2$,故只须 $W\_l-\_l+\frac{\_^2l^2}{2}\leqslant 0$ 即可,或 $(\_l)^2-2W(\_l)\leqslant 0$,由此可求出 l 的具体范围 L.

至此,对于合适的 C 及 T,只要集合 $L\cap[l_1,\infty)$ 不空,其中的元素全部为模型的解.

特别地,当信息到达服从强度为 λ 的 Poisson过程且 Y_1 具有参数为 $\_^{-1}$ 的指数分布时,若 $C_0(l)=C_0$,由(12)知

$$EC_T(t)/t\to C_0\lambda e^{l}/(e^{l}-1).$$

由(3)知上式右侧关于 l 严格递减,此时单位时间上空白信息的费用与总费用之比由(8)决定,但(8)关于 l 严增,于是便有

定理 4 若信息到达服从强度为λ的 Poisson过程,信息长度具有参数为 $\_^{-1}$ 的指数分布,且 $C_0(l)=C_0$,则对给定的 C 及 T,模型的解由不等式组(15)确定.

$$\left.\begin{aligned}&\lambda C_0e^{l}/e^{l-1}\leqslant C,\\&1-e^{l}-1/\_le^{l}\leqslant T,\end{aligned}\right\} \tag{15}$$

而且对合适的 C 与 T,(15)必定有解.

3 一个数值例子及进一步的问题

作为说明,我们对某些特定的 $\_$,λ,C,C_0 及 T 采用模拟的方法,求取(15)中的近似解,数值结果反映在文末的表中.

定理3与定理4均为定理1的特殊情形,其中的特别之处在于 $C_0(l)$ 作为费用函数是一个常值,如果在实际问题中具体情况下 $C_0(l)$ 因 l 变化而变化,则定理1中模型的求解将依赖于 $C_0(l)$ 的具体形式,其解的存在性也将决定于 $C_0(l)$ 的具体性质.对 $C_0(l)$ 为线性函数等具体费用函数的情形有待进一步研究.

附表　一些数值模拟结果

Table　some results from numerical simulation

T	C	C_0	λ	_	l
0.42	2	1	1.260	1	0≤ l≤ 1.2
0.21	2	1	1.667	1	0.4≤ l≤ 0.5
0.10	4	2	0.2	1	0.1≤ l≤ 0.2
0.05	2	0.4	0.2	1	0.04≤ l≤ 0.10

参 考 文 献

1 Ross S M. Stochastic process. New York: John Wiley & sons, 1983

2 Barlow R E, Proschan F. Statistical theory of reliability and life testing Probability models. To Begin with, Silver Spring, MD. 1981

3 邓永录,梁之舜.随机点过程及其应用.北京:科学出版社, 1992. 147- 150

4 Feller W. An introduction to probability theory and its applications, Volume 2, second edition. New York: John Wiley & Sons, 1971

The Optimal Solutions of a Packing Problem in Data Communication

Li Xiaohu　Li Zehui

(Department of Mathematics, Lanzhou University, 730000, Lanzhou, China)

Chen Yuanqing

(Department of Science & Technology, Guangdong Development Bank, 510120, Guangzhou, China)

Abstract　A kind of information packing problem in data communication is dealt with. The information groups arrive according to a point process which is partially finite; the lengths of all groups are identical and independent to the arriving process. For safety and reliability, it is necessary to divide each group into some data pieces, pack and convey them by computer. A probability model is built by which the procedure of the packing can be discussed quantitatively. Stochastic methods are used to determine whether the model has really definite optimal solutions. It is shown that if only suitable total cost in average per time unit and ratio of total cost of additional information to that of useful ones in average per time unit are given, such a model has indeed some intervals as its definite solutions. When the information arrivals are a Poisson process with intensity λ and the length of every data group has exponential distribution with parameter $\_^{-1}$, the optimal packing length of information can be found. Some numerical simulations are performed to demonstrate the effectiveness of such models.

Key words　packing　Poission process　expectation

1991 MR Subject Classification　62E17